data efficiency gain of GPT-SELFIES